

Benefits of Multivariate Statistical Process Control Based on Principal Component Analysis in Solder Paste Printing Process where 100% Automatic Inspection is already Installed

Pedro Delgado¹, Cristina Martins¹, Ana Braga², Cláudia Barros², Isabel Delgado¹, Carlos Marques¹, Paulo Sampaio²

¹ Bosch Car Multimedia Portugal SA, apartado 2458, 4705-970 Braga, Portugal

² ALGORITMI Research Centre, University of Minho, 4800-058 Guimarães, Portugal

Abstract. The process of printing and inspecting solder paste deposits in Printed Circuit Boards (PCB) involves a very large number of variables (more than 30000 can be found in 3D inspection of high density PCBs). State of the art Surface Mount Technology (SMT) production lines rely on 100 % inspection of all paste deposits for each PCB produced. Specification limits for Area, Height, Volume, Offset X and Offset Y have been defined based on detailed and consolidated studies. PCBs with paste deposits failing the defined criteria, are proposed to be rejected.

The study of the variation of the rejected fraction over time, has shown that the process is not always stable and it would benefit from a statistical process control approach.

Statistical process control for 30000 variables is not feasible with a univariate approach. On one side, it is not possible to pay attention to such a high number of Shewhart control charts. On the other side, the very rich information contained in the evolution of the correlation structure would be lost in the case of a univariate approach.

The use of Multivariate Statistical Process Control based on Principal Component Analysis (PCA-MSPC) provides an efficient solution for this problem.

The examples discussed in this paper show that PCA-MSPC in solder paste printing is able to detect and diagnose disturbances in the underlying factors which govern the variation of the process. The early identification of these disturbances can be used to trigger corrective actions before disturbances start to cause defects. The immediate confirmation of effectiveness of the corrective action is a characteristic offered by this method and can be observed in all the examples presented.

Keywords: Multivariate Statistical Process Control, Principal Component Analysis, Solder Paste Inspection, Hotelling's T^2 , Squared Prediction Error, Variable Contributions, Normal Operation Conditions.

1 Introduction

The solder paste printing together with 3D Solder Paste Inspection (3DSPI), constitute a key process in surface mount technology and reflow soldering production lines. Number of defects in the end of production line and reliability of Printed Circuit Boards (PCB) solder joints depend strongly on stable and well centred paste printing process.

In Bosch Car Multimedia Portugal, before 2008, the quality of this process was assured through the use of best practices defined in production guidelines, well trained and experienced team of line operators and process engineers, the use of top class machines and raw material, adequate preventive maintenance and regular machine capability evaluation.

From 2008 until 2017, the process was significantly improved by the introduction of 3DSPI: Area, Height, Volume, Offset X and Offset Y of each paste deposit are measured on all PCBs. It is a 100% inspection performed in-line by an automatic measuring system. Lower Specification Limit (LSL) and Upper Specification Limit (USL) for each type of variable were defined based on detailed studies of short term, long-term, within line and between line variation. Stable production periods of more than eight hours without defects after reflow soldering were taken as a starting point. Differences between paste deposit geometries and raw material types, were considered in the specification.

The introduction of 3DSPI was a key factor for quality improvement and cost reduction in Bosch Car Multimedia Portugal production lines. PCBs with high probability of failure in subsequent process phases, could then be rejected based on objective criteria.

The work presented in this paper was triggered by the following questions: is Multivariate Statistical Process Control based on PCA (PCA-MSPC) framework appropriate for monitoring processes where the number of variables can reach 30000? Will PCA-MSPC control charts help in the identification of assignable causes of variation and contribute to stabilize the process? Can this framework be installed and work efficiently in production lines with cycle times as low as twenty seconds? Is it possible to further improve a process where 100% automatic inspection is already installed?

2 Solder Paste Printing and 3D Solder Paste Inspection

2.1. Process Description

In a SMT production line with reflow soldering, the first important step is paste printing. Solder paste deposits are accurately printed in PCB copper pads where electronic components will later on be placed by high speed and high accuracy pick and place machines. The PCB populated with electronic components placed on top of solder paste deposits, is then submitted to a reflow soldering process with a suitable

temperature profile. Cycle times in SMT production lines can be as low as twenty seconds.

The required volume, position and shape of paste deposits is obtained through the use of a printing stencil with opening holes. An adequate amount of solder paste is placed on the top side of the stencil and pushed by a squeegee device in the direction of the PCB pads as illustrated in Figure 1.

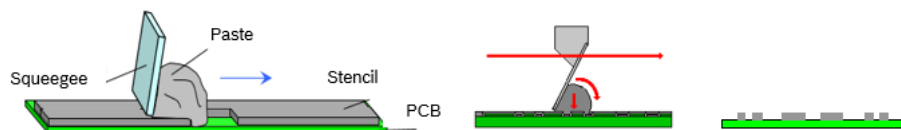


Figure 1: Formation of Solder Deposits.

Most common stencil thicknesses are 150 μm , 125 μm and 112 μm . More than one thickness can be used in the same stencil.

The repeatability of this printing process depends on the correct configuration of different process parameters.

Known root causes of process variation are PCB fixation problems (the PCB must be well supported and stable during the printing process), machine alignment (conveyor, stencil and PCB support must be parallel) PCB warpage, machine printing speed, squeegee pressure, and many others.

Causes of variation are widely described by standardization entities, suppliers of printing machines and solder paste, as well as SMT electronic manufacturers. However, it is essential to know the exact degree in which these causes of variation are affecting Bosch Car Multimedia specific products and production lines.

Area, Height, Volume, Offset X and Offset Y of all deposits of all PCBs are measured by 3DSPI. PCBs with non-conform paste deposits will be rejected. This rejection of faulty PCBs protects subsequent processes and ultimately the client from receiving non-conform products. Another benefit is that historical data becomes available to process engineers who can use this information to evaluate and improve process stability and optimize performance.

2.2. Univariate and Multivariate Statistical Process Control

Even with 100% inspection and correct rejection of non-conform products, process stability is not guaranteed. The analysis of variation in the non-conform rate shows that solder paste printing process still shows some amount of instability. The need of a different kind of process control technique was identified. Detecting negative trends in an early phase and triggering preventive actions before the occurrence of defects, became the new goals (Reis & Gins, 2017; Ferrer, 2014 and Reis & Delgado, 2012).

The well-known features of Statistical Process Control (SPC) based on the distinction of stable versus unstable processes (common causes of variation versus assignable causes of variation) are the classical answer for stabilizing and optimizing processes. This is usually accomplished by monitoring that a variable follows an approximate normal distribution with approximately constant mean and standard

deviation. Shewhart control charts and process capability indexes are the most widely used tools. The classical approach analyses only one variable at a time. Correlations with other variables are not considered. It is immediately recognized that this approach is not feasible when thousands of variables have to be monitored (Montgomery, 2009; Shewhart, 1931).

There are multivariate control charts usable for more than one variable. They monitor the Hotelling's T^2 statistics in order to evaluate the weighted distance of each observation to the centre of the multivariate distribution. The weighting factor is the standard deviation in the direction containing the observation. Difficulties may appear in the use of T^2 control charts as a consequence of the multicollinearity problem: with large number of variables and the existence of strong correlations, the inversion of the correlation matrix is difficult or not possible because it becomes ill-conditioned or singular (MacGregor & Kourti, 1995; Montgomery, 2009).

The chemical industry came up with a solution to this problem using a framework known as Multivariate Statistical Process Control based on PCA also referred to as PCA-MSPC (MacGregor & Kourti, 1995; Nomikos & MacGregor, 1995).

In a first step, the original hyperspace constituted by all the original variables, is rotated in a way that the new variables become aligned with orthogonal directions of maximum variance. The mathematical tool used is the eigenanalysis of the variance-covariance matrix (or the correlation matrix). This method provides a new axis system aligned with the main directions of variation. The transformed variables are ordered by the amount of variation explained. The direction of these new axis is given by the eigenvectors and the variance observed in each one of these directions is given by the eigenvalues. It was observed in many different fields of application that the underlying factors governing the observed variance tends to concentrate in a smaller number of main directions which are then called principal components (PC). The rotation in order to get a new set of orthogonal variables and the dimensionality reduction obtained by retaining only a smaller number of principal components, provides features which makes this framework very attractive for process control (Jackson, 1991; Jolliffe, 2002; Montgomery, 2009; Wold, Geladi, Esbensen, & Ohman, 1987). Some of those features are the following:

- Using a certain number of observations produced under stable conditions, a model can be built which describes the type of variation to be expected if no disturbances happen in the process. The stable period is known as Normal Operation Conditions period (NOC), training set or phase 1.
- Model building and validation, the most computational demanding and time consuming part of this framework, can be made off-line and easily exported to an in-line process (Esbensen, Guyot, Westad, & Houmoller, 2002).
- In order to export the model to the online monitoring engine, it is only necessary to export the mean values, the standard deviation values, the principal component loadings and control limits calculated for the chosen significance level. Mean and standard deviation are used for mean centring and unit variance scaling. Principal component loadings are the coefficients of the linear combinations which performs the PCA rotation and are the key to

compute the scores, that is, the value of the original variables when represented in the new axis system.

- T^2 statistics can easily be calculated in-line by summing the squared value of scores of each observation in each one of the retained principal component new axes.
- T^2 control chart can be used to monitor the distance of each observation to the centre of the model and monitor process stability in phase 2.
- The stability of the correlation structure can also be monitored in phase 2 using a statistics known as Squared Prediction Error (SPE) or Q. A sudden modification of the correlation structure is indicated by a high value of Q.
- The method is reversible: An observation vector in the original space can be transformed to the reduced dimensionality PCA sub-space. An observation represented in the PCA subspace can be converted back to the original variable space with a prediction error which depends on the number of principal components retained. This transformation can be made with simple matrix equations (MacGregor & Kourti, 1995; Martins, 2016).
- A large number of numerical and graphical tools like score scatter plots, score timeline plots, loading plots, T^2 control charts, Q control charts and some others are available.
- When a process disturbance is detected by T^2 or Q control charts, a process is available to compute which original variables have contributed more to this deviation. Intuitive contribution charts are also available.
- Associated with each principal component, it is frequently possible to identify a physical meaning.
- The installation of PCA-MSPC usually leads to early detection of process disturbances, faster diagnostics of root causes of process deviations, increased knowledge about the process and faster validation of effectiveness of corrective actions.

3 Model Building and Real Time Monitoring with PCA-MSPC

The installation of control charts involves two phases. In phase 1, samples are collected which are representative of the full range of acceptable products. Such period of time is usually referred to as Normal Operation Conditions period. Ideally, this period is centred close to nominal values, and the observed variance should be caused only by common causes of variation or other causes of variation which are intrinsic to the process and cannot be completely eliminated. In other words, production should be well centred and stable in NOC period (Montgomery, 2009; Tracy, Young, & Mason, 1992).

Having a good model is a crucial element for an effective process control. In order to obtain a good model, it is necessary to select a sample which is representative of future acceptable production lots, exclude outliers and use cross validation to define the number of principal components to retain. Expertise in both PCA-MSPC and solder

paste printing technologies is required for building models which will work correctly in production monitoring.

PCA models created in the scope of this work usually retain six principal components and explain approximately 50% to 60% of the observed variance as shown in Figure 2.

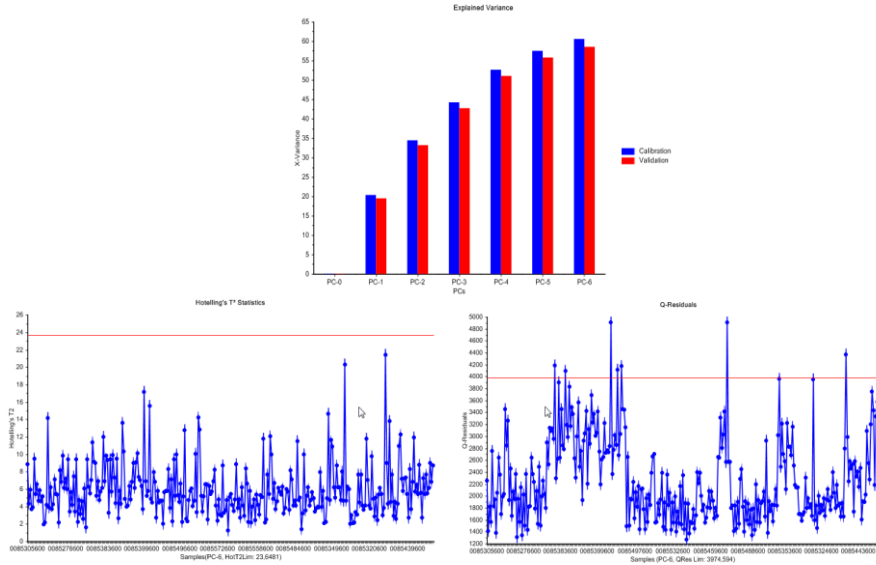


Figure 2: Number of PC retained, explained variance, T^2 and Q residuals control charts in phase 1.

In phase 2, also known as control phase, new observations are measured and compared to the model. The intention of the comparison is to decide if the differences can be explained by common causes of variation or if the differences observed can only be explained assuming the occurrence of an assignable cause of variation (Montgomery, 2009; Tracy, Young, & Mason, 1992).

The first level of this evaluation is made in a cockpit chart with control charts for T^2 , Q and principal component scores. In this work, control limits are calculated for a significance level of 0,1%. If the cockpit chart shows instability, then a deeper dive can be made through the contribution plots in order to identify the original variables affected by the previously detected disturbance.

In this work, the software used is *The Unscrambler X Process Pulse II*® from CAMO Software AS. *The Unscrambler*® is used to create the model (phase 1) and *Process Pulse II*® is used to monitor the process (phase 2).

The described set-up is able to detect the existence of assignable causes of variation like outliers, trends, oscillations or other unusual patterns. T^2 and Q act as summary statistics; timeline principal component scores provide some degree of diagnostic ability since they are frequently associated to a physical meaning; raw data and contribution plots show in detail which original variables contributed to the disturbance. Examples of such cockpit charts are shown in Figure 3 and Figure 4.

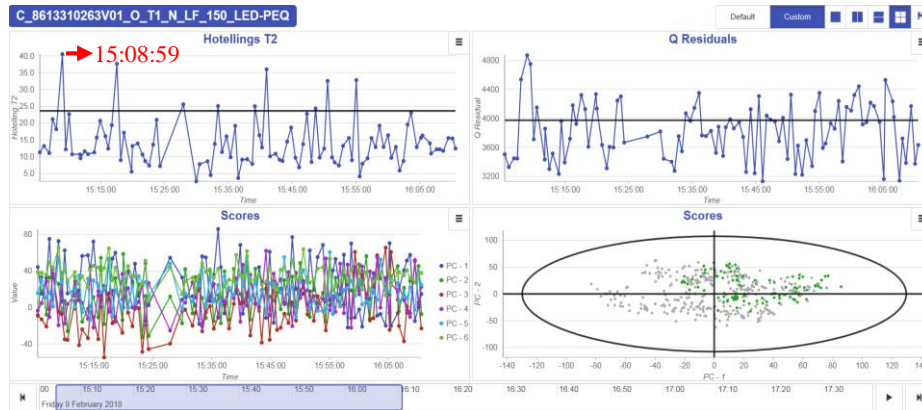


Figure 3: Cockpit Chart using *The Unscrambler X Process Pulse II®*.

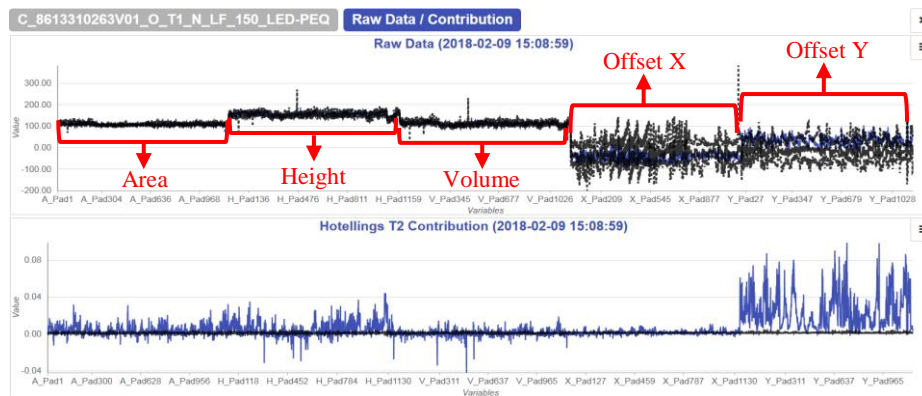


Figure 4: Raw Data and Hotelling's T^2 contribution plots for PCB at 15:08:59.

In raw data and contribution plots (Figure 4), continuous black (or yellow) lines represent the mean values obtained in NOC period. Black (or yellow) dashed lines, represent minimum and maximum values observed in each variable in NOC period. If possible, more than 250 observations are used to build the model. Excluding abnormal circumstances, these black lines are expected to cover a zone of approximately three standard deviations away from the mean value. Blue lines, represent the current observation. The raw data plot can be graphed in the original variable scale or in mean centred and unit variance scale. Area and Volume original units are percent points; Height, Offset X and Y are expressed in μm .

As shown in Figure 3, PCB printed at 15:08:59 has high values of T^2 and Q . Contributions plot and raw data (Figure 4) show high contributions for many Offset Y variables.

Studies made using information collected in NOC period (different production lines/products) have indicated that the physical meaning frequently associated with principal components are:

- PC1 associated to printing direction affecting mainly Offset Y and in a smaller degree Area, Height and Volume.
- PC2 associated with Area, Height and Volume influenced by PCB solder mask thickness, stencil cleaning cycle, machine parameters like squeegee pressure, printing speed, panel snap-off, stencil and squeegee wear-out and many others.
- PC3, PC4, PC5 and PC6, if all retained in the model, are frequently associated to PCB X and Y translations or rotations with different rotation centres.

In some products, physical meaning of principal components can be different but the ones described above are the more frequent.

In the next section, selected case studies illustrate the potential of PCA-MSPC applied to the monitoring of solder paste printing and associated inspection process. The examples presented were collected during six months in four production lines. For these production lines, forty PCA models are already installed.

For each model created, a report is issued containing process parameters used, number of observations, amount of variation explained by the model and checked with cross-validation, T^2 and Q control charts to evaluate the presence of possible outliers, loading plots to illustrate original variable correlations, score scatter plots to evaluate possible existence of clusters, score timeline plots to evaluate process stability. If existing and clearly documented, physical meaning of principal components is included in the report.

4 Results

4.1. Damaged Squeegees

The cockpit chart in Figure 5 shows a production period with high instability due to strong oscillation in T^2 , Q, PC1, PC2 and PC5. T^2 consecutive observations are alternatively inside and outside control limits. This behaviour is typical of problems associated with alternated printing directions.

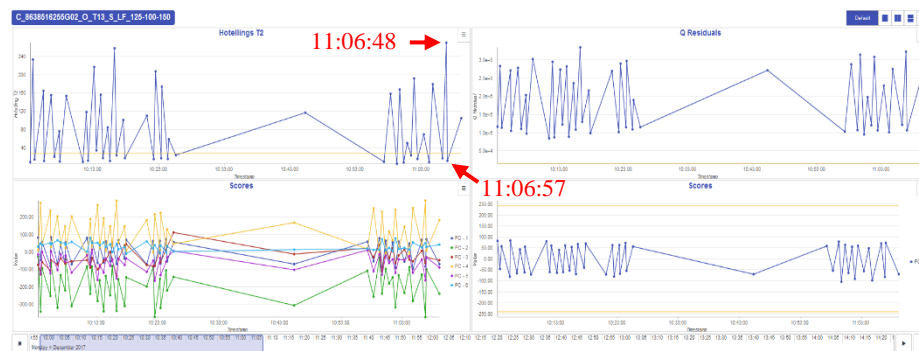


Figure 5: Process instability associated to forward and backward printing direction.

Figure 6 and Figure 7 show raw data and contribution plots for two consecutive observations, the first being outside control limits and the second being inside control limits. The plots in Figure 6 show excessive amount of paste visible in Height and Volume variables. It should be noted that not all Height and Volume variables are affected. The plots in Figure 7 show that all variables are close to model centre.

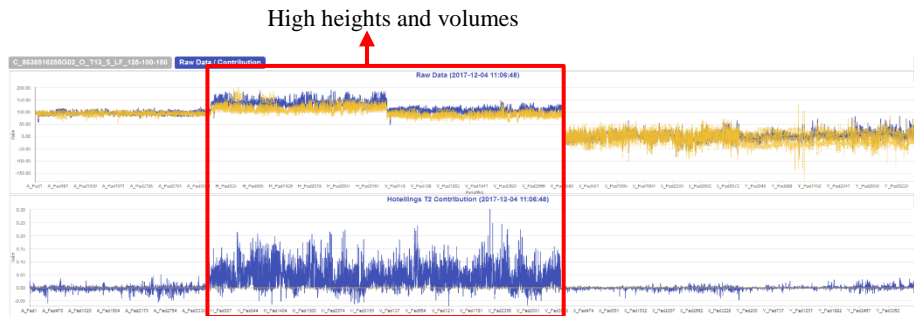


Figure 6: Raw data and contribution plot for PCB 11:06:48.



Figure 7: Raw data and contribution plot for PCB 11:06:57.

When the problem is associated to printing direction, the squeegee is the most probable root cause. The squeegee was replaced at 11:15 and T^2 , Q, PC1, PC2 and PC5 returned to a position closer to the model centre as can be confirmed in Figure 8.

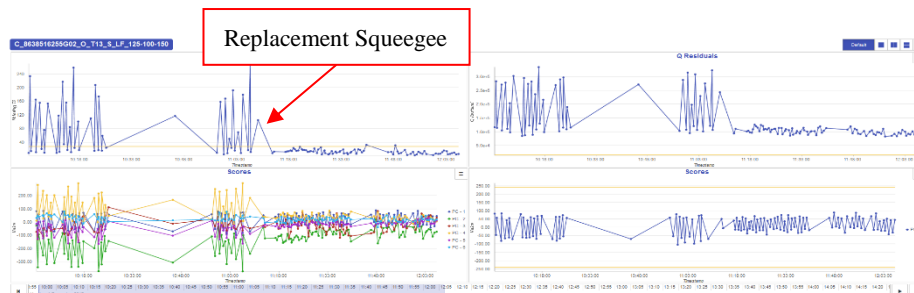


Figure 8: Squeegee replaced at 11:15.

The removed squeegee was inspected with backside illumination and some wear out zones became visible as in Figure 9.

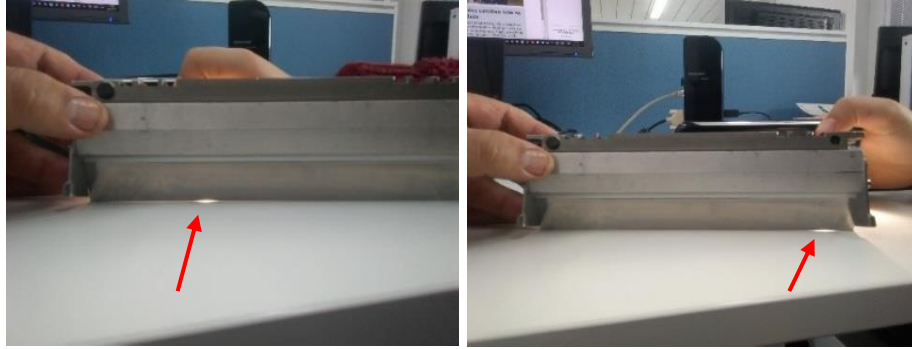


Figure 9: Wear out zones in a damaged squeegee with backside illumination.

Damaged squeegees associated to alternate printing directions is a known problem which frequently appears and remains affecting production quality during long periods of time. Operators and maintenance have been informed that PCA-MSPC is effective in the early detection of this problem.

4.2. Different PCB Suppliers

It is to be expected that PCB coming from different suppliers show different results. These differences, if not too large, are part of the variation which we cannot be avoided.

Figure 10 shows a production process where T^2 and PC1 show some deviation from the centre of the model. At approximately 15:09, T^2 and PC1 changed suddenly approaching the centre showing a stable process. Line operator informed that a new lot of raw PCBs was introduced in the line.

In this particular case, PC1 physical meaning is the difference between supplier 1 and supplier 2. In order to double check this conclusion, the operator was asked to reintroduce in the line PCBs from supplier 1. As expected, T^2 and PC1 returned to the initial condition, as shown in Figure 11.

If the difference from PCB suppliers is not too large, this assigned cause of variation can be accepted and included in the model. If it is too large, the inspection machine measurement program has to be adapted performing an operation called bare board teaching.



Figure 10: T^2 and PC1 sudden change caused by PCBs from a different supplier.

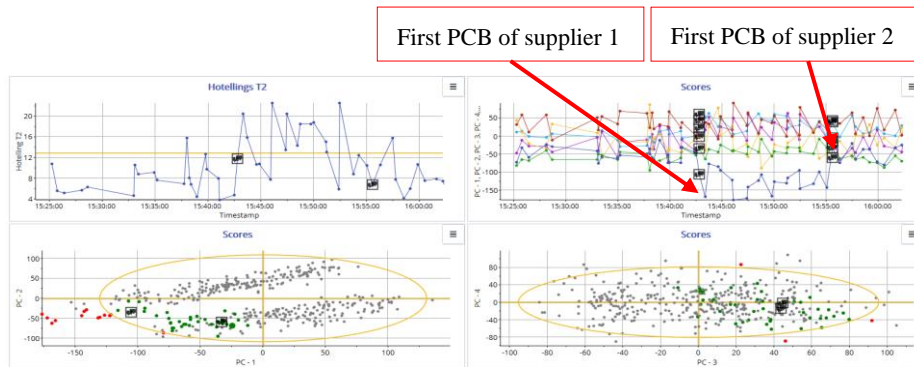


Figure 11: Confirmation of different T^2 and PC1 results caused by different PCB suppliers.

4.3. Impact of Production Line Interruptions of Small to Medium Duration.

In this example, it was observed that the first PCB produced after two-hour production line stoppage shows decreased Volumes, Areas and Heights of paste.

Figure 12 shows stable production until around 06:00. The production line had interruptions until 8:00.

The PCB produced at 08:35:26 shows sudden change in T^2 , Q and PC2 statistics. Raw data and contribution plot in Figure 13 show high contribution from many Volume, Area and Height variables. This reduced amount of paste, was related to dried paste in some stencil openings caused by production line stoppage. The same pattern happened around 9:00 after an interruption of forty minutes.

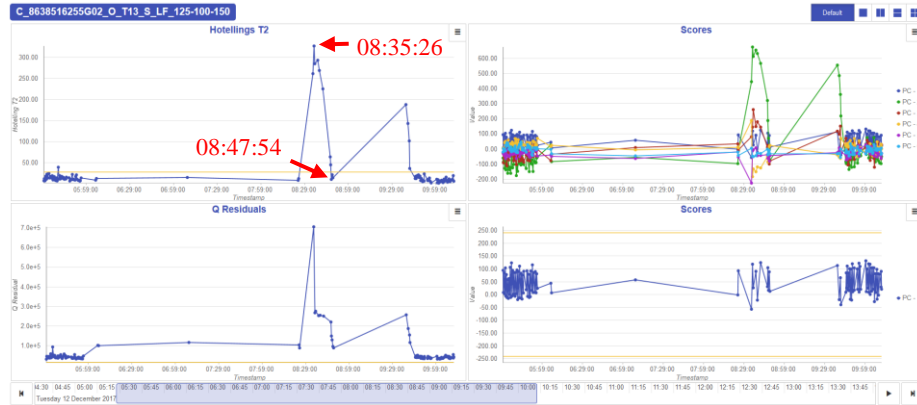


Figure 12: Impact of medium duration line interruption.

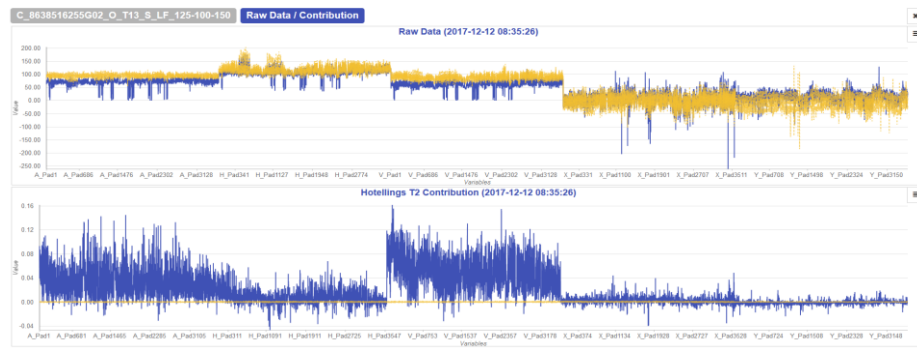


Figure 13: Raw data and contribution plot after a medium duration line interruption. Note 0,16 maximum in contribution plot yy axis.

Figure 14 shows raw data and contributions for the PCB produced at 08:47:54 with values close to the model centre.

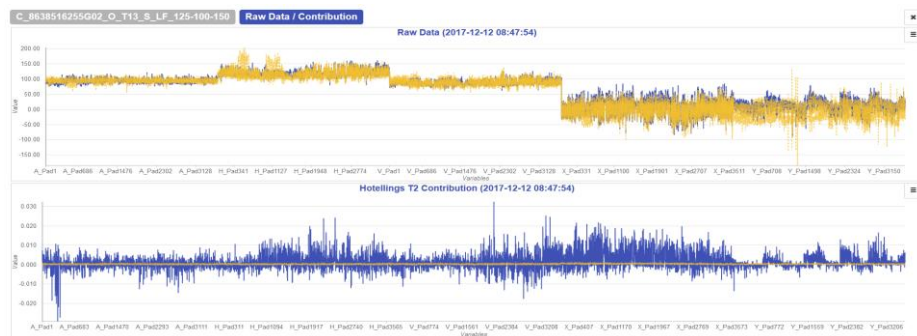


Figure 14: Raw data and contribution plot for an observation close to model centre. Note 0,03 maximum in contribution plot yy axis.

5 Quantification of Improvement

Process monitoring using PCA-MSPC was installed in production line 15 (SMT15) in March 2017. SMT15 was chosen as a pilot line because it runs high density PCBs. The installation of PCA models for products running in the line was made during March and April. Some optimizations were made and the new system was in full operation by the end of May. Process disturbances and their associated root causes were identified and corrected. The evolution of First Pass Yield (FPY) is shown in Figure 15.

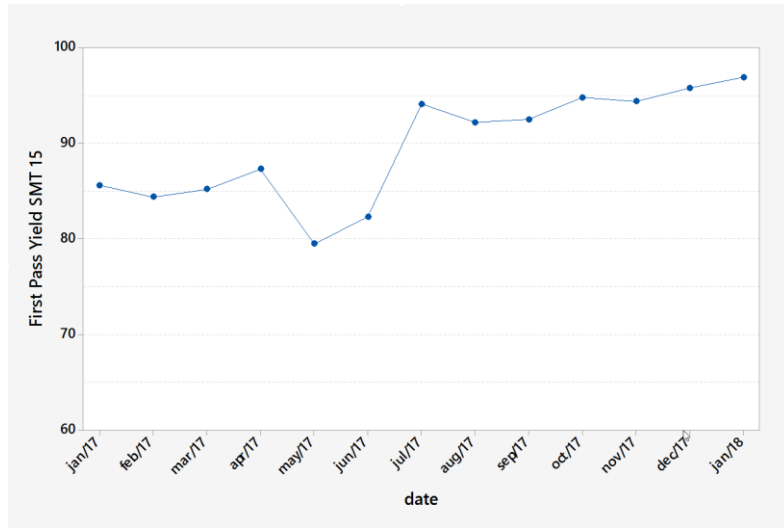


Figure 15: First Pass Yield evolution in SMT15.

Due to the sustained improvement observed, the system has been recently installed in three additional production lines and will be extended to all thirty SMT lines until the end of 2018.

A process engineer specialized in PCA-MSPC and solder paste printing, will be in charge of a centralized performance monitoring. Every disturbance identified will be analysed and maintenance will introduce corrective actions when appropriate. Effectiveness of corrective actions, will also be confirmed immediately after the maintenance intervention.

In order to quantify the results obtained in eight months, a six sigma metrics is used. First Pass Yield (FPY) and Defects Per Million Opportunities (DPMO) are calculated and compared to six-sigma long term quality level of 5.4 DPMO.

The FPY is 0.97 as shown in Figure 15, number of defect opportunities assumed for reference is 5000 (high density PCBs running in this line). The base error rate of the process is expressed in DPMO. The average number of Defects Per Unit (DPU) is $5000 \times \text{DPMO} \times 10^{-6}$. The approximation given by the Poisson distribution, $\text{FPY} = e^{-\text{DPU}}$ provides an estimation of 6 DPMO. This is a very good result close to six sigma quality level of 5.4 DPMO.

6 Conclusions

When the number of variables to be controlled is very high (thousands or tens of thousands), 100% automatic inspection is important because it avoids that defective products reach further steps of the production process or the final client. It is to be underlined that this is a full automatic process made under control of a computer program at high speed and without human intervention. However, 100% automatic inspection is not enough to guarantee stable processes. Assignable causes of variation are frequently present without being identified by process engineers. If these causes are not identified and corrected, the process will drift and the rejection rate and associated costs will rise.

Benefits of using PCA-MSPC applied to solder paste printing have been identified:

- Early detection and replacement of wear-out or damaged squeegees;
- Early detection and replacement of damaged or wear-out stencils;
- Early detection and correction of machine degradation in axis systems, motors, clamping devices, support bases or other machine organs;
- Early detection of mistakes causing wrong parameter adjustments as printing speed, squeegee pressure, PCB snap-off, type and periodicity of cleaning cycle, and others;
- Early detection of differences in solder mask thickness indicating the need to perform bare board teaching;
- More accuracy in the diagnostic of root causes of disturbances;
- Reduction of variation due to over-adjustment caused by wrong diagnostic and consequent inappropriate corrective action;
- Process engineer teams improve their knowledge about the process and become more motivated and capable to improve it.

The introduction of PCA-MSPC in all thirty Bosch Car Multimedia SMT lines is the next step in the direction of better process monitoring, lower costs and improved quality.

Another important conclusion is that PCA-MSPC framework worked well with a number of variables as high as 30000 and a number of six principal components explaining an approximate value of 50% to 60% of total observed variance.

T^2 and principal component score control charts behave according the expectations and show good sensitivity and specificity. Q control charts show frequently stable values but out of the control limits calculated in phase 1.

One way to improve Q specificity, is to build the model using a sample that contains observations taken from lots produced in different days. A model build in this way is more representative of future production, Q statistics behaviour gets better but sensitivity of T^2 and principal component score statistics decreases slightly.

Future work will be done in order to improve Q specificity without degrading T^2 sensitivity.

Acknowledgement

This work is supported by: European Structural and Investment Funds in the FEDER component, through the Operational Competitiveness and Internationalization Programme (COMPETE 2020) Project n° 002814; Funding Reference: POCI-01-0247-FEDER-002814

References

- Esbensen, K. H., Guyot, D., Westad, F., & Houmoller, L. P. (2002). *Multivariate data analysis: in practice: an introduction to multivariate data analysis and experimental design*. Multivariate Data Analysis.
- Ferrer, A. (2014). Latent structures-based multivariate statistical process control: A paradigm shift. *Quality Engineering*, 26(1), 72-91.
- Jackson, J.E. (1991). *A User's Guide to Principal Components*. New York: John Wiley and Sons.
- Jolliffe, I. T. (2002). *Principal Component Analysis* (2nd ed). New York: Springer.
- MacGregor, J. F. & Kourti, T. (1995). Process analysis, monitoring and diagnosis, using multivariate projection methods. *Chemometrics and Intelligent Laboratory Systems*, 28(1), pp. 3-21.
- Martins, C. (2016). *Controlo Estatístico Multivariado do Processo*. University of Minho.
- Montgomery, D. C. (2009). *Introduction to Statistical Quality Control* (6th ed.). New York: John Wiley & Sons.
- Nomikos, P., & MacGregor, J. (1995). Multivariate SPC Charts for Monitoring Batch Processes. *Technometrics*, 31(1), 41-59.
- Reis, M. & Delgado, P. (2012). A large-scale statistical process control approach for the monitoring of electronic devices assemblage. *Computers and Chemical Engineering*, 39, 163-169.
- Reis, M. S., & Gins, G. (2017). Industrial Process Monitoring in the Big Data/Industry 4.0 Era: From Detection, to Diagnosis, to Prognosis. *Processes*, 5(3), 35.
- Shewhart, W.A. (1931). *Economic control of quality of manufactured product*. Vol. Republished in 1980 as a 50th Anniversary Commemorative Reissue by ASQC Quality Press. New York: D. Van Nostrand Company, Inc.
- Tracy, N. D. , J.C. Young, and R.L. Mason (1992). Multivariate control charts for individual observations. *Journal of Quality Technology*, 24(2), 88-95.
- Wold, S., P. Geladi, K. Esbensen, and J. Ohman (1987). Multi-way principal components and PLS analysis. *Journal of Chemometrics*, 1, 41-56.